November 21, 2021

Public Comments on Draft Guidance (FDA-2020-D-2307), "Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products"

We appreciate the opportunity to provide our comments on the FDA's Draft Guidance, FDA-2020-D-2307. Overall, we applaud the agency's effort for its comprehensiveness and thoughtfulness in providing guidance about leveraging real-world data (RWD) to generate real-world evidence (RWE) for drug and biologic products.

As academic physicians who study the use of RWE and leverage it for research about the safety and effectiveness of medical products, we would like to specifically note our support for several key provisions, described below, that are essential to ensuring the rigor of RWD-based studies. While there is significant potential to leverage RWD, there are also many limitations that could lead to inaccurate conclusions; these limitations must be understood and addressed if RWD-based studies can be conducted to inform regulatory decision-making. In many cases, clinical trials (ideally with more pragmatic than explanatory elements) will continue to be needed. We note that this draft guidance, originating from the Center for Drug Evaluation and Research, applies to drug and biologic products; given the increasing focus on real-world data to inform regulatory decision-making, we think that many issues are similarly pertinent and apply to vaccines, gene therapies, regenerative medicines, and medical devices.

We also offer comments, organized by section, that we think can be used to strengthen this important Draft Guidance.

I. Introduction and Scope

- 66-67: We agree that reliability includes accuracy, completeness, provenance, and traceability. In particular, traceability is crucial to ensuring high-quality, unbiased RWD-based research. We further think that FDA should mandate registration and results reporting of all RWD-based studies of FDA-regulated medical products – including prespecification of data sources; prespecification of primary, secondary, safety, and negative control endpoints; prespecificaton of the statistical analysis plan; and public reporting of results for all analyses to coincide either with submission for publication or within 6 months of study completion (Dhruva SS, Shah ND, Ross JS *Mayo Clin Proc* 2020;95:2609-2611). These steps are necessary to ensure the responsible conduct of robust, trustworthy RWD-based research and should apply to all RWD-based research, regardless of sponsor or investigator (including, for example, industry-led/sponsored RWD-based research, government-led/sponsored RWD-based research, and academic-led/sponsored RWD-based research).
- 68-69: We agree with FDA that "sufficient numbers of representative patients" should be included. We encourage FDA to define "representative patients" – including the requirement that patients have a similar demographic profile (age, gender, race, ethnicity) and have similar comorbidities to those who would be indicated to potentially

receive a drug or biologic in real-world clinical practice. Additionally, representative patients should approximately match disease severity and subtype, especially when being considered for post-approval studies or new indications. We think that definitions would be crucial to helping ensure that evidence is generated for populations who are often underrepresented in clinical trials – including older adults, women, racial and ethnic minorities, and those with comorbidities – as this is a key strength of RWD.

III. General Considerations

- 98-99: We fully agree with FDA that "sponsors should submit protocols and statistical analysis plans before conducting the study," as noted above re: lines 66-67 to support the rigor of RWD-based studies. Use of the STaRT-RWE template (Wang SV et al *BMJ* 2021;372:m4856) could be recommended as a very helpful starting place, with the possibility of including submission of additional study design elements, as appropriate.

IV. Data Sources

- 223-225: We agree that data sources should contain adequate numbers of patients with adequate follow-up for outcomes of interest. However, FDA should state specifically that follow-up within an electronic health record (EHR) may be insufficient; there must be sufficient demonstration of detail for visits in which appropriate outcomes may be ascertained. As an example, if the exposure of interest is a drug used for cardiovascular disease and follow-up is at 1-year for safety and effectiveness outcomes, if the patient has follow-up at 1-year documented within the health system's EHR but only within ophthalmology care and not cardiology care, then it is unlikely that this would count as sufficient follow-up. Because patients may go to specialty centers for specific medications or procedural/surgical care and receive the remainder care in a different setting, the issue of follow-up must be carefully addressed. This is addressed in lines 145-150, stating that data captured within a health system may not represent comprehensive care; however, additional detail would be helpful.
- 232-233: In addition to the lack of systematic capture of nonprescription drugs or those that are not reimbursed, FDA should note that these are also often not updated in EHRs; they may be copied and carried forward without updates and, thus, not only missing but also sometimes documented as being taken by a patient long after they are no longer being used.
- Additionally in the issue of medication exposure, FDA should note that EHRs do not necessarily accurately capture medications that are discontinued. Specifically, when a clinician discontinues a medication in the EHR, that medication discontinuation order is rarely transmitted to the pharmacy and the medication may continue being dispensed until refills are exhausted. This is particularly problematic for medications dispensed through mail-order pharmacies and clinicians are thus forced to rely on patients to stop taking medications when discontinued. This quote from a peer-reviewed manuscript describes the issue, "Because most EHRs are not configured to notify pharmacies of prescription cancellations electronically, clinicians and staff are required to call the pharmacy to communicate that a prescription has been discontinued or actively submit a fax communication to the pharmacy about the prescription discontinuation. It is

estimated that up to 5% of prescriptions are still filled after intended discontinuation because of a lack of communication of intent to the pharmacy." (Nelson SD, Kumah-Crystal Y, *JAMA Internal Med* 2021;181:1384-1385). Overall, this means that ideally pharmacy claims data are also used to verify exposure of interest (as noted in lines 680-682) – as well as to verify other covariates that may be medications.

- 276-280: We fully agree with FDA that patients usually have health care data at multiple sites and it is essential to accurately de-duplicate data to address overcounting and, similarly, address the fact that only a fraction of a patient's total health care history will be available at a given health system. Although FDA notes in line 283 that this is not an issue with a multi-site hospital network, there are sometimes differing amounts of retrospective data available (for example, if one hospital in a health system came online to an EHR after other hospital systems – patient data within that specific health system may be more limited). This is very important to investigate and account for prior to commencing any analyses.
- 371-393: We agree with FDA that there are many innovative solutions that are looking to leverage unstructured data, including artificial intelligence. While these have significant potential, we agree that the key factors described in lines 389-393 should be universally available if these tools are used to mine unstructured data.
- 399: We agree with FDA that missing data is critical to address – including the fact that tests or treatments may not have been ordered, ordered but not conducted/received, may have been conducted/received but not captured in the data source, or the data were unavailable. Because health services are usually obtained in clinical practice only if they will change management (as opposed to in a clinical trial, where they are often part of a protocol), there is significant risk for confounding by indication for covariate and outcome ascertainment. These issues must be addressed in any RWD-based study.
- 479-480: In addition to sensitivity and specificity, FDA should make clear that "predictive values" explicitly refer to the PPV and NPV. Additionally, the accuracy should be evaluated.

V. Study Design Elements

- 530-532: We fully agree with FDA about the requirement to address temporal changes in standard of care, availability of other treatments, diagnosis criteria, and other study design factors. Older data may not reflect standards of practice for a multitude of reasons, and these should always be acknowledged and addressed.
- 545-550: We agree with FDA that key variables should be validated. ICD-9, ICD-10, and CPT codes represent valuable sources of data but there are various factors that affect coding and they may not be reflective of the intended conceptual definition. Validation is a time- and resource-intensive process but is critical because administrative claims are used for billing purposes – not for conducting RWD-base studies (which is a secondary use of these codes).
- 617-624: We agree with FDA about the importance of uncaptured prescriptions, which occur commonly such as through generic drug discount programs or other out-of-pocket purchases.

- 645-646: FDA makes excellent points that patients may not refill prescriptions on time or may refill them early. Although FDA notes (superscript 13) that this guidance does not address issues related to medication adherence, we think that this should at least be mentioned. Medication adherence is crucial to determining exposure because intended/prescribed use may differ from the actual use in clinical practice. Patients may not take medications as prescribed for a larger variety of reasons - some of which are related to adverse effects of the medications. Reasons for non-adherence, including side effects, should also be ascertained in RWD-based studies and added to this Guidance.
- 709: One important additional consideration that should be included is the use of falsification endpoints / negative controls. Falsification endpoints, which are different from the exposures being tested, are not causally related to the intervention and, therefore, should not be affected by the intervention. Falsification endpoints should be pre-specified and, if they are not found to be positive associations, would support the overall conclusions of interest. (Prasad V and Jena AB, *JAMA* 2013;309:241-242).
- 767-771: We agree with FDA that outcomes with well-defined diagnostic criteria are more likely to be consistently captured in RWD than those that are subjective in nature. It is often difficult or impossible to identify key study design elements with administrative claims or EHR data as they are unlikely to appear in a structured form. Published literature on that topic shows that only 15% of clinical trials published in high-impact medical journals could be feasibly replicated using administrative claims or EHR data (Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS *JAMA Netw Open* 2019; e1912869). A similar limitation was identified among medications approved in 2011 (Fralick M, Bartsch E, Darrow JJ, Kesselheim AS *Pharmacoepimiol Drug Saf* 2020; 29:1273-1278). Additionally, there is still some heterogeneity to the determination of events such as stroke and myocardial infarction. For example, recent peer-reviewed evidence indicates that the introduction of ICD-10 codes for types 2 through 5 myocardial infarction were associated with a decrease in hospitalizations for ICD-10 codes that were suggestive of type 1 myocardial infarction MI (McCarthy CP et al *J Am Coll Cardiol* 2021;78:1242-1253). In this section, FDA also refers to proxy measures or multi-component definitions that may need exploration and justification for use (lines 766-767); it would be helpful if FDA can explicitly define what validated proxy measures are acceptable.
- 784-788: FDA notes that patient-generated data should be defined, constructed, validated, and collected. These are all important points for patient-generated data (including patient-reported outcomes or PROs). FDA should add to the guidance information about pre-specifying in greater detail how PROs will be collected because the inability to ascertain patient-reported outcomes from real-world data remains a critical limitation since these are rarely integrated into electronic health records and never in claims (Basch E *N Engl J Med* 2017;376:105-108).
- 834: We agree that a standardized process is important to ensure consistent application of criteria to validate outcomes. Additionally, we recommend that, when feasible, such validation efforts are initially performed in a non-analytic dataset. This will allow refinement of methods prior to final outcome validation and reduce risk of bias.

- 838-843: We agree with FDA about the importance of addressing observer bias. Observers should not be individuals involved in the care of the patient. Additionally, redacting identifiable data about exposure status in EHR data is challenging and time-consuming, but is ideal to ensure rigor.
- 885: We agree with FDA that measuring PPV alone is often inadequate. We recommend that FDA explicitly state the NPV should always be assessed and reported in an appropriate sample of cases.
- 956-957: In addition to sensitivity, specificity, PPV, and NPV, we recommend that FDA include "accuracy" in the list of indices that are reported for quantitative bias analysis.
- 997-1003: We agree with FDA about the many unmeasured or imperfectly measured confounders. We also recommend that FDA note that even though there are codes for items such as alcohol use disorder, they are unlikely to be comprehensive and alternative means are necessary for accurate ascertainment.

VI. Data Quality During Data Accrual, Curation, and Transformation into the Final Study-Specific Dataset

- 1073: In addition to the specification of data provenance, we recommend that FDA require datasets be submitted to FDA for independent analysis and validation.
- 1101-1102: In addition to continuity of coverage, we recommend that FDA require comparisons of patients with and without continuous coverage to ensure that patients not included in the study do not differ significantly from those who were included.
- 1107-1120: We agree that updates or changes to coding practices are important to include. In particular, the transition from ICD-9 to ICD-10 has created a data chasm (Khera R, Dorsey KB, Krumholz HM *JAMA* 2018;320:133-134).
- 1243-1248: We agree with FDA that all submitted programs should be thoroughly annotated to describe the data management and analyses and to enable FDA replication.

Sincerely,

Sanket Dhruva, MD, MHS
University of California, San Francisco

Reshma Ramachandran, MD, MPP
Yale School of Medicine

Joseph S. Ross, MD, MHS
Yale University